மின்சொற்களஞ்சியம் தயாரித்தல்

முனைவர் ச. இராஜேந்திரன்
மொழியியல் துறை
தமிழ்பல்கலைக்கழகம்
தஞ்சாவூர்

அறிமுகம்

தற்காலத்தமிழுக்கு ஒரு மின்சொற்களஞ்சியம் (electronic thesaurus) உருவாக்க வேண்டும் என்ற எண்ணத்தோடு இவ்வாய்வு மேற்கொள்ளப்பட்டது. ஜோன்ஸ் (1986:201) சொற்களஞ்சியம் (thesaurus) என்பதற்குப் பின்வரும் பொருளில் விவரணை தருகின்றார்:

"ஒரு சொற்களஞ்சியத்தில் ஒரு பொருள்பன்மொழிகள் இருக்க வேண்டிய கட்டாயம் இல்லை. மற்றும் ஒரு ஒருபொருள்பன்மொழி அகராதி ஒரு சொற்களஞ்சியமாக அமையத்தேவையில்லை; விரிவான பொருளில் சொற்களஞ்சியம் என்பது சொற்களைக் கருத்துருக்கள் (concepts), தலைப்புகள் (topics) அல்லது பாடப்பொருள்கள் (subjects) ஆகியவற்றின் அடிப்படையில் பாகுபாடு செய்வதாகும்; ஒரு வகுப்பில் (class) தரப்பட்டுள்ள சில சொற்கள் ஒருபொருள் பன்மொழிகளாகும் என்ற செய்தி நிச்சயமாகத் தேவையற்ற ஒன்றாகும். மாறாக ஒருபொருள் பன்மொழி அகராதி ராஜெஸ்டின் சொற்களஞ்சியத்தில் (Roget's Thesaurus) தரப்பட்டுள்ள கருத்துருப் பாகுபாட்டைத் (conceptual classification) தராமல் சில குழுமங்களாக வரும் சொற்களைத்தான் பகுத்தளிக்கும். இது ஒரு பக்குவமற்ற வேற்றுமைப்படுத்தலாகும்; சொற்களஞ்சியத்தில் ஒருபொருள்பன்மொழிகள் அடிக்கடி காணப்படும்; ஒருபொருள் பன்மொழி அகராதிகளில் உள்ள குழுமங்கள் பெரிதாகவும் ஒருபொருள்பன்மொழியத்தின் பரந்த விளக்கத்திற்கு மாதிரியாகவும் அமைந்துள்ள குழுமங்களை இணைக்கும் குறுக்கு நோக்கீட்டுகளையும் கொண்டிருக்குமானால் அவை சொற்களஞ்சித் தன்மை பெறலாம்."

சொற்களஞ்சியமும் அகராதியும்

சொற்களை அகராவரிசையில் ஒருங்குபடுத்தி சொல் தொடர்பான செய்திகளைத் தரும் மரபுமுறை, ஒரே பொருள் தரும் சொற்களை ஒன்றுசேர்த்தும் ஒன்றுக்கொன்று தொடர்புள்ள அல்லது ஒரே பொருள் தரும் சொற்களை அங்குமிங்குமாகச் சிதறித் தரும். ஆனால் துரதிருஷ்டவசமாக வெளிப்படையான வேறு வழி இல்லாமையாலும் மேற்சொன்ன மரபுவழி பழக்கத்தில் இருந்து வந்தது. அகரவரிசைச் சொல் வரிசை மூலம் செய்திகளைப் பெறுவடு சிரமமானதாகவும் கால விரையம் மிகுந்ததாகவும் இருந்தது. பெரும்பான்மையோர் அகராதியைப் புரட்டி வேண்டிய செய்டிகளைப் பெறுவது தங்கள் வேலையைப் பாதிக்கும் என்ற காரணத்தால் அகராதியை விரும்புவதில்லை.

இன்றைய கணினி யுகத்தில் இதற்குத் தீர்வு கிடைத்துள்ளது. மின் அகராதி அல்லது கணினி அகராதி உருவாக்குவதற்கு முக்கியமான காரணம் யாதெனில், மின் அகராதியிலிருந்து மிக விரைவாகத் தகவல்களைப் பெற இயலும். வேண்டிய சொல்லைச் சாவிப்பலகையில் தட்டியோ நேரடியாகத் தேர்ந்தெடுத்தோ அகராதித் தகவல்களைப் பெறலாம். பெரும்பாலும் அகராதிகள் கணினியால் அச்சடிக்கப் பெறுவதால் இந்த மின் அகராதிகளைப் பொருத்தமான சொல்தரவு அடிப்படைகளாக (database) எளிதில் மாற்ற முடியும். இம்மின் அகராதியைக் கொண்டு தேவையான சொல் பற்றிய விளக்கங்களை நொடிப் பொழுதில் பெறலாம். அகராதியில் பதிவு செய்துள்ள தகவல்களை மீளப்பெற்றுத்தரும் கணினியின் அதிவேக மீள் தகவல் பெறும் சக்தி இங்கு உதவுகின்றது. இதுபோல் கணிப்பொறியில் தகவகல்களை ஆயும் திறன், ஒன்றிணைக்கும் திறன் போன்ற எண்ணற்ற திறன்கள் உள்ளன. இத்திறன்களையெல்லாம் மின் அகராதியைக் கணிப்பொறி வழி உருவாக்கவும் நோக்குவதற்கும் பயன்படுத்துகின்ற பொழுது மின் அகராதி நல்ல நோக்கீட்டுக் கருவியாக அமையும். நாம் இங்கு முன்மொழியும் மின்சொற்களஞ்சியம் மரபுவழி அகராதிச் செய்திகளையும் பொருட்புல அகராதிச் செய்திகளையும் இணைத்து மிக விரைவாகச் செயல்புரியும் கணினியைப் பயன்படுத்தி உருவாக்கும் சக்திவாய்ந்த கருவியாக அமையும்.

இருபதாம் நூற்றாண்டில் உருவான உளமொழியியல் மொழி அறிதிறனைப் புலனறிவு அடிப்படையில் ஆய்ந்தது. மொழியியலாரும் உளமொழியியலாரும் மொழியறிவின் அமைப்பையும் சொல்லறிவின் அடிப்படையையும் ஆய்ந்தனர். மில்லர் மற்றும் ஜான்சன் லைர்ட் என்பார் சொல்லறிவு ஆய்வு உளச் சொல்லியல் ஆய்வாக கருடப்படவேண்டும் என்று கூறினர். அன்றாடச் மொழிச்செய்திகளை உருவாக்கவும் புரிந்துகொள்ளவும் ஒலியனியல், தொடரியல் சொல்லியல் என்பன சேர்ந்து செயல்பட வேண்டும் என்பதில் மொழியியலாளர்கள் தெள்ளத்தெளிவாக உள்ளனர். உளமொழியியலாளார் மன அகராதியின் ஒரு கால மொழியியல் அடிப்படையிலான கூறுகளைக் கண்டுபிடித்துள்ளனர். தொடக்க காலத்தில் சொல்தகவல் அடிப்படையை உருவாக்குவதற்கான முயற்சிகள் எடுத்துக் கொள்ளப்பட்டன. சொற்களை அகரவரிசை அடிப்படையில் அல்லாமல் கருத்துக்கள் அடிப்படையில் தேட முயற்சிகள் எடுத்துக் கொள்ளப்பட்டன. இதற்கு மரபு அடிப்படையில் அமைந்த ஒரு மின் அகராதி துணைநின்றது. பின்னர் சொல் பற்றிய தகவல்களின் பரிணாமம் கூடக்கூட சொல்வலை (word net) என்ற அளவில் இத்தகைய செயல்முறை விரிந்துள்ளது.

உளமொழியல் கோட்பாடுகளை எந்த அளவிற்குப் பயன்படுத்தி மின்சொற்களஞ்சியம் அமைக்கலாம் என்பது ஆய்வுக்குரியது. பெரும்பாலும் இத்தகைய ஆய்வுகள் பெயர்களைச் சுற்றி வருகின்றனவே தவிர பிற சொல்வகைப்பாட்டைச் சார்ந்த வினைகள், பெயரடைகள் பற்றிய ஆய்வுகள் அதிகம் இல்லை. இச்சொற்களஞ்சியம் ஒரு அகராதியை அடிப்படையாகக் கொண்டு அதில் உள்ள சொற்களைப் பொருள் அடிப்படையிலும் அவற்றின் பொருண்மை உறவு அடிப்படையிலும் தொடர்புபடுத்தும் நோக்கில் அமைந்துள்ளது. சொற்களஞ்சியத்தைப் பொறுத்தமட்டில் பெயர், வினை, பெயரடை, வினையடை ஆகிய சொல்வகைகளைச் சார்ந்த சொற்கள் தரவுகளாகத் தேர்ந்தெடுக்கப்பட்டுள்ளன. இலக்கணச் சொற்கள் தொடர்பன்கள் என்ற நிலையில் சேர்த்துக் கொள்ளப்பட்டுள்ளன. தேர்ந்தெடுக்கப்பட்ட சொற்கள் யாவும் பொருண்மைக்கூறுகள் அடிப்படையில் வகை படுத்தப்பட்டு வகைப்பாட்டு எண்களுடன் பட்டியலிடப்பட்டு தரவு அடிப்படையாகச் சேகரிக்கப்பட்டுள்ளன.

மின்சொற்களஞ்சியத்தின் மிகப் பரந்த நோக்கம் என்னவென்றால் சொற்தகவல்களை சொல்வடிவுகள் அடிப்படையில் அல்லாமல் சொற்பொருண்மை அடிப்படையில் ஒழுங்குபடுத்துவதாகும். இந்த அடிப்படையில் மின்சொற்களஞ்சியம் அகராதியிலிருந்தும் வேறுபடும். ஆங்கிலத்திற்கு உருவாக்கப்பட்ட ரொஜஸ்ட் தெசாரஸ் போன்ற சொற்களஞ்சியங்களும் தமிழுக்கு இராசேந்திரனால் (இராசேந்திரன், 2001) உருவாக்கப்பட்ட தற்காலத் தமிழ்ச் சொற்களஞ்சியமும் தமிழ் மின்சொற்களஞ்சிய உருவாக்கத்திற்கு முன்னோடியாய் அமைகின்றன. மேற்கூறிய சொற்களஞ்சியங்கள் புத்தக வடிவில் அமைந்தவை. அகரவரிசையில் அமைந்த சொற்களஞ்சியங்களின் சிக்கல் என்னவென்றால் அவற்றில் தேவைக்கு அதிகமான/மிகையான பதிவுகள் காணப் பெறும். இரு சொற்கள் ஒரு பொருள்பன்மொழிகள் என்றால் கூட அகரவரிசைக்காக இருமுறை பதிவு பெறும். தலைப்பு ஒழுங்குமுறையில் அமைந்த சொற்களஞ்சியங்களின் சிக்கல் என்னவென்றால் இரண்டுமுறை நோக்கீடு செய்துதான் தேவையான தகவல்களைப் பெரும்பாலும் பெற இயலும். முதல்முறை அகரவரிசையில் அமைந்த சொல்லடைவு மூலம் குறிப்பிட்ட சொல் எந்தப் பொருண்மைத் தலைப்பில் வருகிறது என்பதைத் தெரிந்து கொண்டு இரண்டாவது முறை நோக்கீடு செய்துதான் வேண்டிய தகவலைப் பெறமுடிகிறது. இதனால் தேடும் நேரம் இரண்டு மடங்காக உயரும். கணிப்பொறியின் வழி தேடலை விரைவாகச் செய்ய இயலும்.

சொல்வடிவுகளுக்கும் சொல் பொருண்மைகளுக்கும் உள்ள பொருத்தம் பல. சில வடிவுகள் வேறுபட்ட பொருண்மைகளைத் தரும்: சில பொருண்மைகளைப் பல வேறுபட்ட வடிவங்களில் வெளிப்படுத்தலாம். பல்பொருள் ஒருமொழியம் (polysemy), ஒருபொருள் பன்மொழியம் (synonymy) என்பன அகராதியியலின் இரு கடினமான சிக்கல்களாகும். இவற்றைப் பொருத்துவதின் துணைநிலைகளாகக் கொள்ளலாம். உள அகராதியை (mental dictionary) அணுகுவதன் வாயிலாக பல்பொருள் ஒருமொழியமும் ஒருபொருள் பன்மொழியமும் சிக்கல்களாகத் தெரிகின்றன. ஒரு வடிவத்தைத் தெரிந்து கொள்ளும் கேட்பவரோ படிப்பரோ இந்தப் பல்பொருள் ஒருமொழியத்தை எதிர்கொள்ள வேண்டி வரும்; பொருளை வெளிப்படுத்த விரும்பும் பேசுபவர்

அல்லது எழுதுபவர் ஒருபொருள்பன்மொழிகளுக்கிடையில் ஒன்றை தேர்ந்தெடுக்க வேண்டிவரும்.

மின்சொற்களஞ்சியம் பொருண்மை உறவுகளுக்கும் சொல் உறவுகளுக்கும் இடையே உள்ள வேறுபாட்டை உணர்த்துகின்றது. பொருண்மைகளுக்கு இடையே உள்ள வேறுபாட்டை உணர்த்துகின்றது. பொருண்மைகளுக்கு இடையே உள்ள பொருண்மை உறவுகளுக்கு முக்கியத்துவம் கொடுக்கப்பட்டாலும், சொற்களுக்கு இடையே உள்ள உறவுகளும் உட்படுத்தப்பட்டுள்ளன.

சொல் ஆய்வியில் மின் சொற்களஞ்சியத்தின் பங்கு

சொல் ஆய்விகள் (word processors) (புத்தகம், கட்டுரைகள் போன்ற) உரைகளை உள்ளீடு செய்து சேகரிக்கவும் பல்வேறு பட்ட எழுத்துருக்களில் உரைகளை அமைத்து அச்சிட்டுக்கொள்ளவும் பயன்படுகின்றது. சொல் ஆய்விகள் செய்திகளைக் கணிப்பொறியில் உள்ளீடு செய்து வேண்டியபடி அதில் உள்ள நுட்பங்களைப் பயன்படுத்தி வடிவமைத்துக்கொள்ளப் பயன்படுகின்றன. ஆங்கிலத்திற்கென்று WODSTAR, MS WORD, WORD PERFECT, STAR WORD போன்ற பல சொல் ஆய்விகள் உள்ளன. தமிழுக்கும் பல சொல் ஆய்விகள் உருவாக்கப்பட்டுப் பயன்படுத்தப்படுகின்றன. பொதுவாகச் சொல் ஆய்விகளில் உரைகளை நகல் எடுக்கவும் வெட்டி இணைக்கவும் நீக்கம் செய்யவும் பல வடிவங்களில் தகவலை மாற்றி அமைக்கவும் தேவையான வசதிகளைப் பதிவுப் பட்டியல் தருகின்றது. கருவிப்பட்டியலில் உள்ள வசதிகள் கீழே பட்டியலிடப்பட்டுள்ளன.

1. எழுத்து மற்றும் இலக்கணத் திருத்தம் செய்யும் வசதி
2. சொற்களைத் தேர்ந்தெடுக்கப் பயன்படும் சொற்களஞ்சியம் (Thesaurus)

கருவிப் பட்டியலிலுள்ள சொற்களஞ்சியத்தைப் பயன்படுத்துபவர் உரையிலுள்ள ஒரு சொல்லுக்குப் பதிலாக வேறு ஒரு சொல்லைத்

தேர்ந்தெடுக்கவும் அச்சொல்லால் பதிலீடு செய்யவும் இயலும். சொற்களஞ்சியத்தைச் சுட்டியால் அழைத்தால் அது உரையிலுள்ள சொல்லுக்கு ஈடான ஒருபொருள்பன்மொழிகளையும் எதிர்மொழிகளையும் தொடர்புடைய சொற்களையும் தரும். எடுத்துக்காட்டாக MS WORD குறைந்தது பின்வரும் சொற்களஞ்சியச் செய்திகளைத் தரும்.

1. உரையில் வரும் சொல்லைப் பதிலீடு செய்வதற்கு ஒருபொருள்பன்மொழிகளின் பட்டியல்.
2. சொல்லின் பொருள்
3. தொடர்பான சொல் அல்லது சொற்கள்

எடுத்துக்காட்டாக, MS WORD-ஐப் பயன்படுத்துவோர் electronic என்ற சொல்லுக்கான தகவலைச் சொற்களஞ்சியத்தில் தேடினால் உரையாடல் பெட்டி பினவரும் தகவலைக்காட்டும்.

Replaceable Synonyms

automatic, automated, mechanical, computerized, programmed, streamlined, cybernetic

Meaning
automatic (adj.)

Related item
electron

நாம் electron என்பதற்கான தகவலைச் சொற்களஞ்சியத்தில் தேடினால் பின்வரும் மூவகைத் தகவல் கிடைக்கும்.

Replaceable Synonyms

particle in an atom, atom, negative particle, negation, electrically charged element, neutron, proton, particles

Meaning

particle of an atom(noun)

Related item

electron

நாம் இங்கு முன்மொழியும் மின்சொற்களஞ்சியம் ஒரு அகராதியையும் (dictionary lexicon) ஒரு பொருட்புல அகராதியையும் (thesaurus) மிஞ்சும் வகையில் அமையும். எடுத்துக்காட்டாக ஒருவர் மருத்துவ நிலையத்தைப் பற்றி கட்டுரையோ கதையோ எழுதவேண்டுமானால் அவருக்கு மருத்துவநிலையம் தொடர்பான எந்தச் சொல் தேவைப்பட்டாலும் சொற்களஞ்சியம் தரவேண்டும் என்பது மின்சொற்களஞ்சியத்தின் பரந்த நோக்கமாகும். இது சொற்களின் ஒரு தரவுமூலமாக அமையும் (lexical resource). சுருக்கமாகச் சொல்லப்போனால் சொல்வலை (wordnet) என்பதில் எதிர்பார்கிற எல்லாச் செய்திகளும் இதில் அடக்கப்படும்.

மின்சொற்களஞ்சியத்தின் செயற்பாடு

மின்சொற்களஞ்சியத்தில் பயன்படுத்துபவர்க்கு என்று அமைக்கப்பட்டுள்ள இடைமுகத்தைப் (inteface) பயன்படுத்தி வேண்டிய செய்திகளைப் பெறலாம். கீழே தரப்பட்டுள்ளவை ஒரு சில பயன்பாடுகளாகும். இவை மின்சொற்களஞ்சியத்தின் முதல் நிலை அடிப்படையில் அமைந்தவை.

ஒருபொருள்பன்மொழிகளை அறிந்துகொள்ளுதல்

இடைமுகத்தில் தெரிந்த ஒரு சொல்லைத் தந்து அதனோடு நெருங்கிய தொடர்புள்ள சொற்களைப் பெறலாம். எடுத்துக்காட்டாக *நல்லவர்* என்பதை

இடைமுகத்தில் உள்ளீடு செய்து அது தொடர்பான சொற்களைத் தரச் சொன்னால் *நல்லவர்* என்பதன் ஒருபொருள்பன்மொழிகள் 'பண்புத்தொடர்புடைய பெயர்கள்' என்ற தலைப்பின் கீழ்வரும் *நல்லவர்* என்ற துணைத் தலைப்பின் கீழ் பட்டியலிடப்படும்:

நல்லவர் (நல்லவன், நல்லவள், நல்லவள்), உத்தமர் (உத்தமன், உத்தமி), ஏந்தல்.

பகுதி ஒருபொருள்பன்மொழிகள் மற்றும் தொடர்புடைய சொற்களை அறிந்துகொள்ளுதல்

நல்லவர் என்பதுடன் தொடர்புடைய கூடுதல் சொற்களைத் தரச்சொன்னால் பின்வரும் சொற்களும் பட்டியலிடப்படும்.

பெரியவர் (பெரியவன்), பெரியமனிதர் (பெரியமனிதன்), பெரியோர், உயர்ந்தவர் (உயர்ந்தவன், உயர்ந்தவள்), சிறந்தோர் (சிறந்தோன், சிறந்தோள்), அடிகள், ஆன்றோர், மேன்மக்கள், உத்தமபுருஷர், மகாபுருஷர், மகான், மகாத்மா, மானஸ்தர் (மானஸ்தன், மானஸ்தி), மானி

எதிர்மறைப் பொருளில் வரும் சொற்களை அறிந்து கொள்ளுதல்

'பண்பு தொடர்புடைய பெயர்கள்' என்ற தலைப்பின் கீழ் *நல்லவர்* என்ற சொல்லுக்கு எதிரிடையான சொற்களையும் அறிந்துகொள்ளலாம்.

கெட்டவர் (கெட்டவன்:கெட்டவள்), கொடியவன்:கொடியவள்), கொடியோர் (கொடியோன்:கொடியோள்), தீயவர் (தீயவன்:தீயவள்), பொல்லாதவர் (பொல்லாதவன்: பொல்லாதவள்), குருரர் (குருரன்:குரூரி), வக்கிரர் (வக்கிரன்), கிராதகர் (கிராதகன்:கிராதகி), துஷ்டர் (துஷ்டன்:துஷடை), போக்கிரி, வம்பர் (வம்பன்:வம்பி), விஷமி, அக்கிரமி, அயோக்கியர் (அயோக்கியன்), கயவாளி, கசவாளி.

படிநிலை அமைப்பின் மூலம் உள்ளடங்கு-உள்ளடங்கு பொருண்மைத் தொடர்பு மற்றும் பொருண்மைக் கூறுகளை அறிந்து கொள்ளுதல்

நல்லவர் என்பது படிநிலை அமைப்பில் பின்வருமாறு அமைவதைப் பார்த்தோம்:

பருமைப்பெயர்கள்>உயர்திணைப்பெயர்கள்>மனிதர்கள்>பண்புகுறித்த பெயர்கள்>நல்லவர் மேற்கண்ட படிநிலை அமைப்பிலிருந்து நல்லவர் மேற்கண்ட படிநிலை அமைப்பிலிருந்து *நல்லவர்* என்ற சொல் பண்புகுறித்த பெயர்கள் என்பதன் உள்ளடங்குச் சொல்லாகவும் பண்புகுறித்த பெயர்கள் யாவும் மனிதர்கள் என்பதன் உள்ளடங்குச் சொற்களாகவும் மனிதர்கள் என்பதன் கீழ்வரும் சொற்கள் யாவும் உயர்திணைப்பெயர்கள் என்பதன் உள்ளடக்கு சொற்களாகவும் உயர்திணைப்பெயர்கள் பருமைப்பெயர்களின் உள்ளடக்குச் சொற்களாகவும் வருவதை அறிந்துகொள்ளலாம். மட்டுமன்றி அவை உள்ளடக்கும் பொருண்மைப் பண்புக்கூறுகளையும் அறிந்து கொள்ளலாம்.

நல்லவர்
+ பருமைப்பெயர்
+ உயர்திணைப்பெயர்
+ மனிதப்பெயர்
+ பண்புகுறித்தபெயர்

சொற்கள் பொருண்மைக்கூறுகளால் ஒன்றுபடலையும் வேறுபடலையும் அறிந்து கொள்ளுதல்

நல்லவர் என்பதன் ஒருபொருள்பன்மொழியான *உத்தமர்* என்ற சொல்லுக்கும் *நல்லவர்* என்பதற்குத் தரப்பட்ட பொருண்மைக்கூறுகளைத் தரலாம்.

உத்தமர்
+ பருமைப்பெயர்

+ உயர்திணைப்பெயர்
+ மனிதப்பெயர்
+ பண்புகுறித்தபெயர்

எதிர்மறைப்பொருளில் வரும் *கெட்டவர்* என்ற சொல் *நல்லவர்* என்ற சொல்லிலிருந்து 'பண்பு' என்ற பொருண்மைக்கூறால் வேறுபடலையும் உணரலாம்.

கெட்டவர்
+ பருமைப்பெயர்
+ உயர்திணைப்பெயர்
+ மனிதப்பெயர்
+ பண்புகுறித்தபெயர்

இங்கும் இதற்கு முன்னர் கூறிய பத்திகளில் உள்ள செய்திகள் ஒரு அகராதிச் சொற்களஞ்சியம் (thesaurus with alphabetical entries) உருவாக்கப்பயன்படும்.

சொற்களின் விளக்கப்பொருளை அறிந்துகொள்ளுதல்

சொற்கள் சொற்களஞ்சியத்தில் வரும் இடம் அதன் பொருளைத் தீர்மானிப்பதோடு அச்சொற்களின் பொருள் விவரணையையும் தருகின்றது. எடுத்துக்காட்டாக *நல்லவர்* என்ற சொல் பண்மைக்குறிக்கிற மனிதப் பெயர் என்ற பொருள் பெறும். *அறிஞன்* என்ற சொல்லுக்கு அறிவுத்தொடர்புடைய மனிதப் பெயர் என்ற விளக்கத்தைப் படிநிலை அமைப்பிலிருந்து பெறலாம். *கொடு* என்ற வினைக்கு ஒருவரின் கையிலிருந்து ஒரு பொருள் மற்றொருவர் கைக்குப் போவதைக் குறிக்கும் உடைமைமாற்ற வினை என்ற விளக்கம் பெறலாம். *திங்கள்* என்பது வாரத்தின் முதல்நாள் என்ற விளக்கத்தைப் பெறலாம். *கண்* தலையின் பகுதி என்ற விளக்கத்தைப் பெறலாம்.

மின்சொற்களஞ்சிய உருவாக்கம்

முன்னுரை

சொற்களஞ்சியம் என்று அழைக்கப்படுகின்ற பொருட்புல அகராதி அதன் விரிந்த தற்காலப் பொருளில் சொற்களை கருத்துருக்கள், தலைப்புகள், அல்லது பாடப்பொருள்கள் அடிப்படையில் பாகுபாடு செய்வதாகும். எடுத்துக்கொள்ளப்பட்ட மின் சொற்களஞ்சியம் இயற்கைமொழி ஆய்வால் தரப்பட்ட அஸ்திவாரத்தின் மேல் கட்டப்பட்டுள்ளது. மின் சொற்களஞ்சியம் உருவாக்குகையில் நேரிட வேண்டிய சில மொழியியல் வாதப்பொருட்களும் பின்பற்ற வேண்டிய சில கணினி வழிமுறைகளும் உள்ளன. பின்வருவன மொழியியல் வாதப்பொருள்களாகும்.

1. சொற்களுக்கிடையில் உள்ள சொல் உறவுகளை நிறுவுதல்
2. குறைந்தபட்சம் ஒத்தல், உள்ளடங்கல், இணக்கம், இணக்கமின்மை என்ற நான்கு சொல் உறவுகளால் ஆன படிநிலை வகைப்பாட்டை நிறுவுதல் மற்றும் வேறுபடுத்துகிற பொருண்மைப் பண்புக்கூறுகளைக் கொண்டு சொற்களங்களை நிறுவுதல்.
3. சொற்களைத் தேர்ந்தெடுத்தல்
4. சொற்களை இறுதி சொற் களத்தில் வரிசைப்படுத்தல்
5. அடிப்படைப் பொருளை ஆக்கப்பொருளுடன் தொடர்புபடுத்தல்

சொற்களஞ்சிய உருவாக்கத்தை இருவகையில் செய்யலாம். ஒன்று கீழிருந்து மேல்நிலைக்கு வருவது. அதாவது சொற்களுக்கிடையே உள்ள உறவை முதலில் நிறுவிவிட்டு அதிலிருந்து சொற்களை சொல்வகுப்புகளாகவும் சொற்களஞ்சியமாகவும் பகுப்பது. இதை ஆங்கிலத்தில் bottom-up method என்று குறிப்பிடுவர். இது நல்ல முயற்சி என்றாலும் கால தாமதம் ஏற்படும். மற்றொரு வழி மேலிலிருந்து கீழ்நிலைக்கு வருவது. இதை ஆங்கிலத்தில் top-down method என்று கூறுவர். இந்த வழிமுறை தான் இங்கு கையாளப்பட்டுள்ளது.

ஏற்கனவே இராசேந்திரனின் (2001) அச்சிட்ட தற்காலச் தமிழ்ச் சொற்களஞ்சியம் இருப்பதன் காரணமாக இரண்டாவது முறை எளிதாக அமைந்தது.

தரவு அடிப்படை (Database)

இராசேந்திரானில் (2001) தரப்பட்ட தமிழ்சொற்றொகையின் பொருண்மை அமைப்பு (அதாவது படிநிலைப் பாகுபாட்டு ஒழுங்கமைப்பு) வழியமைப்புமுறைகளால் (programs) மின்சொற்களஞ்சியத்திற்கான தரவு அடிப்படையாக மாற்றப்பட்டுள்ளது. இதிலிருந்துதான் பயன்படுத்துபவர் கேட்கும் கேள்விகளுக்கான தகவல் தரப்படும். சொற்களஞ்சியச் சொற்றொகை பின்வருமாறு பொருண்மைப்புலங்களாகப்பகுக்கப்பட்டுள்ளது (semantic fieds), ஒவ்வொரு பொருட்புலனுக்கும் படிநிலை அடிப்படையிலும் முன்நிலை அடிப்படையிலும் எண்கள் தரப்பட்டுள்ளன. இவ்வெண்களின் வரிசை சொற்றொகைப் பாகுபாட்டின் பொருண்மை அமைப்பை விளக்குவதாய் அமையும். தரவு அடிப்படை புதிய சொற்களைச் சேர்க்கவும் புதிய பாகுபாடுகளை உருவாக்கவும் வேண்டி நெகிழ்வுள்ளதாய் அமைக்கப்பட்டுள்ளது. இரண்டாம் நிலையில் தரப்படும் புதிய செய்திகள் தற்போதைய தரவுடிப்படையுடன் உறவுபடுத்தப்படும். பின்வரும் சொற்றொகைப் பாகுபாடு மின்சொற்களஞ்சியத்தின் முதன்நிலை தரவு அடிப்படையை உருப்படுத்தம் செய்யும்.

பயன்படுத்துவோருக்கான இடைமுகப்பு

பயன்படுத்துவோருக்கான இடைமுகப்பு வழி மின்சொற்களஞ்சியத்தைப் பயன்படுத்துபவர் முன்நாம் குறிப்பிட்ட சொல்குறித்த பல செய்திகளை அறிந்து கொள்ளலாம்.

நிறைவுரை

மின்சொற்களஞ்சிய உருவாக்கத்தின் பரந்தநோக்கங்களையும் அது எந்த அளவுக்கு அதன் முதன் நிலையில் நிறைவேற்றப்பட்டுள்ளன என்பது குறித்தும் தற்போதைய நிலையில் மின்சொற்களஞ்சியத்தைப் பயன்படுத்திப் பெற இயலும் செய்திகள் குறித்தும் பார்த்தோம்.

* * * * *